



Papeles el tiempo de los derechos

SOBRE EL CONTENIDO DE LA ROBOÉTICA

Rafael de Asís

Instituto de Derechos Humanos Bartolomé de las Casas
Departamento de Derecho Internacional, Eclesiástico y Filosofía del Derecho
Universidad Carlos III de Madrid

Palabras Clave: Roboética, Derechos humanos, Tecnoética, Agencia moral, Robótica.

Número: 10

Año: 2013

ISSN: 1989-8797

Comité Evaluador de los Working Papers “El Tiempo de los Derechos”

María José Añón (Universidad de Valencia)
María del Carmen Barranco (Universidad Carlos III)
María José Bernuz (Universidad de Zaragoza)
Manuel Calvo García (Universidad de Zaragoza)
Rafael de Asís (Universidad Carlos III)
Eusebio Fernández (Universidad Carlos III)
Andrés García Inda (Universidad de Zaragoza)
Cristina García Pascual (Universidad de Valencia)
Isabel Garrido (Universidad de Alcalá)
María José González Ordovás (Universidad de Zaragoza)
Jesús Ignacio Martínez García (Universidad of Cantabria)
Antonio E Pérez Luño (Universidad de Sevilla)
Miguel Revenga (Universidad de Cádiz)
Maria Eugenia Rodríguez Palop (Universidad Carlos III)
Eduardo Ruiz Vieytez (Universidad de Deusto)
Jaume Saura (Instituto de Derechos Humanos de Cataluña)



Reconocimiento – NoComercial – SinObraDerivada (by-nc-nd)

No se permite un uso comercial de la obra original ni la generación de obras derivadas.

SOBRE EL CONTENIDO DE LA ROBOÉTICA

Rafael de Asís

Instituto de Derechos Humanos Bartolomé de las Casas

Departamento de Derecho Internacional, Eclesiástico y Filosofía del Derecho

Universidad Carlos III de Madrid

1.- Introducción

Desde el enfoque de los derechos, los problemas éticos surgen cuando se adoptan decisiones que pueden proyectarse en los derechos de otros o en los derechos de uno mismo. La primera de las proyecciones es la menos cuestionada en cuanto problema ético; la segunda, puede serlo algo más. En efecto, difícilmente discutimos la consideración como problema ético de situaciones en las que la actuación de alguien interfiere en los derechos de otros. Sin embargo, cuando se trata de autolimitaciones o de autointerferencias, discutimos si se trata de un problema ético o no.

En ambos casos resulta esencial la determinación de quienes son los otros o de quien es uno mismo; esto es, quienes son los agentes de discurso ético. Así, la inclusión de la robótica en este discurso obliga a tener en cuenta la existencia de distintos tipos de robots y a determinar su papel dentro de este discurso.

Existen diferentes clasificaciones de los robots y de las máquinas. Si lo que se busca es analizar su proyección en la ética, una buena clasificación consistiría en diferenciar entre máquinas dependientes y máquinas independientes. Las primeras son aquellas que requieren de un ser humano para llevar a cabo su función; las independientes son las que no lo necesitan (al menos por regla general). Ciertamente, la existencia de robots completamente independientes es todavía una ilusión. Sin embargo se trata de una distinción que nos sirve para delimitar los problemas al poner de manifiesto dos de los grandes referentes de la ética: la autonomía y la responsabilidad.

Otro camino interesante consiste en utilizar un criterio de clasificación referido a la interacción del robot con los humanos y sus derechos; esto es, un criterio que tiene en cuenta los posibles problemas éticos que plantea el tipo de robot. En esta línea, es posible diferenciar tres tipos de robots (todos ellos estrechamente relacionados): (i) robots que hacen una tarea o un servicio que puede plantear un problema ético por el tipo de servicio o por la relación con el ser humano; (ii) robots que hacen un servicio

que plantea un problema ético ya sea por el daño a terceros, por el posible daño a la propia persona o por el posible daño a la sociedad; (iii) robots que presentan problemas de agencia moral.

Entre los primeros es posible diferenciar tres grupos. Por un lado están los robots industriales manipuladores cuyo principal problema ético radica en su acción sobre el mercado de trabajo. Por otro están aquellos robots que realizan un servicio que plantea problemas éticos en la sociedad como pueden ser los llamados sexbots (robots de sexo). Por último aquellos que realizan una función que tiene incidencia en la atención al ser humano (robots cuidadores o de asistencia).

Entre los segundos es posible diferenciar también entre tres grupos. Por un lado aquellos cuyo cuestionamiento ético se produce porque dañan a terceros (como por ejemplo los drones militares). Por otro aquellos que se cuestionan porque se afirma que pueden producir un daño a la propia persona, ya sea a su integridad física o a su integridad moral o identidad (robots prótesis o robots que estimulan el cerebro). Por último estarían los robots que se afirma pueden producir un daño a la sociedad presente o futura (en este grupo podrían estar algunos de los tipos de robots que he señalado antes, si bien, podríamos también pensar en los llamados nanorobots).

Los que he llamado robots que plantean problemas de agencia moral son aquellos respecto a los cuales puede no estar claro quién es el responsable de su actuación ni su estatus moral. Se trata de robots a los que algunos consideran como agentes o sujetos morales, similares a los seres humanos. Dentro de este grupo se encuentran los robots que aprenden, que razonan de forma relativamente independiente, que se autoalimentan, que son capaces de exteriorizar sentimientos (*Learning, reasoning, eating, o sentimental robots*).

Aunque las clasificaciones anteriores pueden matizarse, nos permiten entrever los principales problemas éticos que se presentan como contenido de la roboética. En este sentido es posible destacar los diez siguientes problemas: (i) Afectación de mercado de trabajo con la consiguiente pérdida de puestos de trabajo; (ii) Integridad física y moral de las personas (básicamente la intervención en los derechos de otros); (iii) Identidad humana física (la cuestión de las prótesis), psicológica (posible dependencia para con el robot) o social (aislamientos social y abandono de tareas sociales); (iv) Diversidad cultural y biológica (uniformidad social y contaminación); (v) Privacidad/intimidad (tratamiento y alteración de los datos personales); (vi) Igualdad y no discriminación (igual oportunidad de acceso al conocimiento y no discriminación,

con especial atención a los de grupos vulnerables; (vii) Propiedad (la regulación de las patentes); (viii) Responsabilidad y seguridad por el engaño, defecto o manipulación, (ix) Identificación y registros de los robots; (x) Agencia o papel moral del robot.

Este catálogo puede estructurarse de diferentes maneras. Bostrom y Yudkowsky aluden a dos grandes cuestiones: (i) el papel de los robots en la ética de los seres humanos; (ii) el estatus moral de los robots. Seguiré en términos generales esta diferenciación distinguiendo entre los usos y límites de la robótica y el trato al robot.

Prácticamente, los nueve primeros problemas antes señalados pertenecen al primero de los asuntos, mientras que el último se correspondería con lo que denomino trato a robot. En principio, el campo singular de la roboética sería propiamente el segundo, ya que el primero se refiere a cuestiones que están presentes en otros ámbitos tecnológicos. Sin embargo, en el primero, la roboética también presenta alguna singularidad derivada de la relación entre seres humanos y máquinas (aunque resulta difícil pensar que se puedan establecer relaciones entre humanos y robots similares a las que se producen entre humanos, esto es algo que no hay que descartar de plano).

2.- Los fines de la robótica

Antes del examen de estas cuestiones, conviene que nos planteemos cuales son los fines de la robótica. Como he señalado en “Sobre el marco de la roboética”, para ello podemos basarnos en las normas sobre bioética y biodiversidad. Desde estas cobran sentido tres afirmaciones generales y programáticas:

- 1.- Los avances científicos y tecnológicos deben servir al bienestar de la humanidad, al desarrollo sostenible de todos los países, a la paz mundial y a la protección y conservación de la naturaleza.
- 2.- El interés y el bienestar humano que prevalece sobre el interés exclusivo de la sociedad o de la ciencia, implica la promoción del respeto de la dignidad humana y la protección de los derechos humanos (entre los que se encuentran la vida y la integridad moral, la intimidad, el libre consentimiento, la elección de una forma de vida, el trabajo...)
- 3.- El respeto de la dignidad humana supone respetar la diversidad humana, el medio ambiente, la biosfera y la biodiversidad (obligando a la utilización sostenible de sus componentes y a la participación justa y equitativa en los beneficios que se deriven de la utilización de los recursos).

Pero existe una cuestión que condiciona todo nuestro discurso, que debe presidir la roboética y que se puede expresar a través de las dos preguntas siguientes: ¿por qué queremos robots?; ¿qué tipos de robots queremos? Se trata de dos preguntas cuya contestación condiciona el contenido de esta reflexión. Son dos preguntas complejas y que nos trasladan a grandes problemas éticos.

En todo caso, aunque sea de forma mínima, podemos avanzar algo en su contestación. Creo que todos estaremos de acuerdo en afirmar que queremos robots para ayudar a los seres humanos y a su entorno. Así, buscamos robots que nos presten servicios y que corrijan diversidades no queridas o enfermedades. El logro de este tipo de máquinas puede tener una clara justificación si bien siempre y cuando se respete la diversidad humana y la biodiversidad y tenga un sentido principalmente terapéutico. Otra puede ocurrir si estamos ante técnicas de mejora, siendo conscientes de que el concepto de mejora es, además, problemático.

Pero no buscamos robots para que nos sustituyan ni, en ese sentido, para que se comporten y sean como humanos. Ciertamente, esto último incluso puede ser discutido, si bien exige tener en cuenta qué es lo que consideramos como humano. Se trata de una cuestión que puede ser resuelta desde diferentes ángulos, todos ellos problemáticos, y que incluso puede sufrir modificaciones al hilo de los avances tecnológicos.

En el ámbito de las ciencias humanas, se suele manejar una concepción del ser humano construida desde referentes éticos y estéticos. Los primeros, los éticos, basados en cuatro grandes principios: la capacidad, la autonomía, la independencia y la responsabilidad. Los segundos, los estéticos, basados en una forma o figura definitoria de lo humano (que tiene su manifestación más grotesca en las alusiones que se hacían en algunos Ordenamientos jurídicos a la exigencia de tener “figura humana” para la consideración de persona). La determinación de la idea de ser humano basada en un patrón de capacidad ajeno a la diversidad y de carácter presuntamente universal y abstracto, puede llevar a apartar de forma explícita de la consideración de personas a determinados seres humanos, o a manejar una visión de persona exclusivamente biológica. Así, para algunos, el concepto de persona se reconduce a la posibilidad de ser consciente del futuro y poseer deseos relacionados con él. O también afirmamos que “somos humanos en la medida en que descendemos de seres humanos”. En cierto sentido, la consideración de lo humano, como categoría, es algo que corresponde determinar a los humanos y debe hacerse dentro del marco ético al que aludí en “Sobre

el marco de la roboética”, manejando así una visión abierta y plural, respetuosa con la diversidad y centrada en la naturaleza.

En todo caso, la consideración de lo humano implica también una serie de circunstancias, rasgos y caracteres que forman parte de nuestra identidad tales como la muerte, la dependencia, el dolor, el sufrimiento, el esfuerzo, la imperfección, las limitaciones..., rasgos que probablemente no queramos replicar en el robot, pero que son, en muchos casos, la expresión más evidente de la diversidad humana y de la libertad. Así, en realidad no se estaría produciendo algo humano sino suprahumano, ante lo cual, la pregunta del para qué cobra aun un relieve mayor.

Por otro lado la construcción de una máquina idéntica a un humano además de no estar justificado en relación con determinados tipos de robots¹, y además de no tenerse claro cuál es su finalidad ni sus consecuencias, podemos pensar que es algo que se enfrenta a las normas bioéticas actuales que prohíben por ejemplo la clonación de seres humanos. Y este argumento puede ser extensible a las prótesis y neuroprótesis que podrían relacionarse con la prohibición de comercio con los órganos del cuerpo, prohibición discutida hoy en día (al considerarse que es algo habitual en ciertos campos). Ciertamente, en este punto, cobra un especial relieve la reflexión sobre la integridad y la identidad humana.

3.- Uso y límites de la robótica

La cuestión del uso y los límites de la robótica puede considerarse como tradicional en el campo de la ética ya que se refiere a los límites de la investigación científica y a la responsabilidad de los agentes y usuarios de la tecnología. Se trata de problemas que tienen que ver con la autonomía y la responsabilidad y sobre los que existen parámetros positivos para su tratamiento: los textos normativos de bioética. No obstante, como he apuntado, existe una dimensión nueva en este campo que viene representada por la proyección social y relacional del robot.

Tradicionalmente, se suelen señalar una serie de principios que deben guiar la investigación teórica y aplicada: la proporcionalidad y la ponderación (evaluación de riesgos y ventajas); el rigor y la prudencia; la igualdad en el acceso a la información y la

¹ Por ejemplo, muchos robots de servicio perderían algunas de sus ventajas si tuvieran emociones.

transparencia; el debate plural y amplio; y la exigencia de formación y calificación científica de los científicos y tecnólogos.

Más allá de estos principios, el análisis de las cuestiones roboéticas es en buena parte casuístico y depende del tipo de robots. Atendiendo a la tipología antes descrita, es posible destacar los siguientes cuatro grandes problemas:

En primer lugar la cuestión sobre la responsabilidad. En este punto es posible diferenciar tres posibles agentes de responsabilidad: el creador (científico o tecnólogo), el comerciante y el usuario. Todos ellos, en algunos casos de manera clara y en otros de forma más alejada, son responsables ya sea de forma directa o indirecta. Y cuatro son los aspectos sobre lo que se proyecta esta responsabilidad: finalidad, engaño, defecto o manipulación. Esta responsabilidad que debe tener una proyección tanto actual como futura, tiene que evaluarse y graduarse teniendo en cuenta el tipo de robots.

En segundo lugar la posible alteración de la identidad humana, ya sea en términos físicos, psicológicos o sociales. Los dos primeros casos se pueden producir en relación con máquinas o robots que funcionan como prótesis, pero también es posible situar aquí a los robots de cuidado o asistentes en relación con los cuales puede producirse situaciones de dependencia (además de lo que puede suponer el que en una relación típicamente humana intervengan máquinas) o los sexbots. La sustitución de personas por robots en el cuidado de los niños o de los mayores, no representa sólo problemas de seguridad y responsabilidad por el comportamiento del robot sino, lo que es más importante, un posible aislamiento social o un cambio en la comprensión de las relaciones sociales junto con un más que cuestionable abandono de conductas y obligaciones que tradicionalmente nos identificaban.

En tercer lugar la posible alteración de la diversidad cultural y biológica y con ello del medio ambiente. Situaciones que se pueden dar con los nanorobots o como consecuencia del depósito de los residuos o de las máquinas que ya no funcionan. Pero también con la uniformidad de los robots y con la desaparición de rasgos que han venido caracterizando a algunas personas.

En cuarto lugar, la posible realización de actos contrarios a los derechos. El ejemplo tipo es el de los drones militares que ponen de manifiesto la posibilidad de que estas máquinas se fabriquen para matar seres humanos. También podrían situarse aquí a los sexbots y su uso para la prostitución, si bien, en realidad, la posibilidad de que esta se lleve a cabo por robots, hace desaparecer el cuestionamiento de esta práctica en términos de derechos. E igualmente se debería situar en este punto la posible

intervención en el mercado de trabajo, que afecta sobre todo a los robots industriales. Pertenecen también a este grupo de problemas la creación de robots que aumentan las diferencias de oportunidades entre los seres humanos y la posible transgresión de la privacidad o la intimidad de las personas (que se produce por ejemplo con los robots asistenciales a los que hay que suministrar información precisa sobre la persona a la que asisten).

4.- El trato al robot

Como ya he señalado, el avance en la robótica con la creación de humanoides capaces de aprender, razonar, sentir, ha provocado que desde algunas direcciones se hable de la posibilidad de considerarlos como verdaderos agentes morales e incluso como titulares de derechos.

Hasta ahora, los robots han sido siempre tratados como objetos y no como sujetos. Y ello con independencia de la función que cumplen que, si tiene relevancia moral, a lo sumo permitiría hablar de las máquinas como objetos morales y de la existencia de obligaciones de cuidado hacia ellos. Sin embargo, lo que demandan estas posiciones es la consideración de las máquinas como sujetos.

Si no utilizamos un razonamiento de corte religioso, la consideración de los robots como sujetos morales exigiría que reunieran los rasgos que exigimos a éstos. Ya he señalado como la definición de sujeto moral asociada a ser humano resulta difícil de concretar. Por otro lado, en la actualidad estamos todavía lejos de lograr una máquina que pueda considerarse idéntica a un humano. Además, también he cuestionado que dirijamos nuestra investigación hacia ese objetivo que, incluso, podríamos considerar como algo prohibido por las actuales normas bioéticas.

Sin embargo, quedan abiertas todavía dos cuestiones de diferente significado. La primera de ellas consistiría en plantearnos ¿qué pasaría si se llegaran a crear máquinas idénticas a los humanos? Ciertamente, podemos pensar que se trata de ciencia ficción y que esto nunca será posible. Así, por mucho que se avance en patrones estéticos, gestos, movimientos, es difícil pensar en que se puedan integrar emociones, sentimientos, reacciones hormonales, inteligencia o conciencia. Sin embargo, los avances en el campo de la neurociencia y de las ciencias de la computación nos obligan a no dar por cerrada de manera completa esta posibilidad. En la historia se ha escrito mucho sobre lo que no llegarían nunca a hacer los robots y luego, muchas de esas cosas han logrado hacerlas.

Como es sabido, el acercamiento del robot al ser humano dio un avance espectacular con el nacimiento de la inteligencia artificial (capacidad de razonar de un agente no vivo) a mediados del siglo pasado. La inteligencia artificial buscó en un primer momento la acumulación del mayor número de conocimientos en la máquina. Pero siempre quedaba por conseguir la posibilidad de que la máquina tuviera conciencia. Y es aquí donde entraron los avances en la neurociencia y el estudio de las emociones, desde el presupuesto de que es posible estudiar la emoción al igual de como se ha estudiado la razón. La mente humana no sólo expresa la capacidad de pensar, sino también la de sentir. Así, a mediados de los 90 del siglo pasado, A. Damasio publica su libro *El error de Descartes*, en el que defiende que si un individuo no tiene capacidad de sentir emociones ha perdido buena parte de su habilidad para razonar. Y a partir de este momento no sólo se busca la acumulación de conocimientos en el robot sino también una mayor relación con el mundo real. De esta forma se considera ya la posibilidad de fabricar robots que reconozcan, comprendan y expresen emociones. No obstante, estas emociones serán diferentes a las humanas. Así, detrás de esta posibilidad no está el fabricar “humanos” sino más bien mejorar la relación de los robots con las personas.

En todo caso, es importante tener en cuenta que para considerar al robot como un agente moral habría que configurar la posibilidad de crear robots que pudieran realizar comportamientos inmorales de forma consciente; y eso no se hace y está claro que no debería hacerse. Se ha afirmado que considerar a los robots como agentes morales implicaría el que estos fueran conscientes de las consecuencias de sus acciones y el que fueran capaces de incumplir reglas. Pero claro, surge entonces una pregunta que ya nos habíamos planteado: ¿por qué queremos fabricar máquinas que incumplan nuestras reglas?

Más allá de lo anterior, si se llegaran a crear máquinas idénticas a los humanos, difícilmente podríamos negar su consideración como agentes morales. Así se ha propuesto utilizar el llamado Test de Turing para afirmar la agencia moral del robot, afirmando que si el robot llega a ser equivalente al ser humano y este es agente moral, el robot también debe serlo². Ninguna máquina ha satisfecho de manera completa y general este test.

² El test de Turing es una prueba que se utiliza en el campo de la inteligencia artificial y que viene a decir que si una máquina se comporta en todos los aspectos como inteligente, entonces debe ser inteligente.

Los diferentes argumentos que se pueden utilizar para negar esta agencia moral, pueden ser contestados. Así, ya en 1950 Turing se refería a una serie de objeciones lanzadas a la posibilidad de que los robots fuera inteligentes, destacando que todas ellas podían ser superadas: (i) La objeción teológica: El alma ha sido dada por Dios a los hombres, pero también puede dársele a las máquinas; (ii) La objeción de los que “esconden la cabeza bajo el ala”: Sería horrible que esto ocurriera, pero no parece que sea un argumento fuerte; (iii) La objeción matemática: las máquinas tienen sus limitaciones, pero también los seres humanos; (iv) El argumento de la conciencia: las máquinas no pueden tener conciencia de lo que hacen ni sentimientos (ya me he referido a esta cuestión); (v) El argumento de las distintas capacidades: hay cosas que aparentemente las máquinas no pueden hacer, pero esto es también predicable de los seres humanos; (vi) El argumento del comportamiento informal: no existen reglas que determinen completamente el comportamiento humano, pero sería posible programar de esta forma a una máquina.

Pero es que además, hay que ser consciente de que negar el estatus moral a los robots por la manera en la que se consigue que tengan capacidad de sentir dolor y sufrir no es aceptable ya que se podría utilizar también con seres humanos. Incluso se podría argumentar a favor de esta consideración apelando a la diversidad humana y a su dignidad.

No obstante, como vengo reiterando no buscamos hacer robots idénticos a los seres humanos y esto hace que no podemos extender la agencia moral de los humanos los robots ni integrarlos en la ética de los humanos. Buscamos la perfección en los robots y los humanos no lo somos y, además, la ética de los humanos cobra sentido ante la posibilidad de lo inmoral y no pretendemos hacer robots inmorales.

La otra cuestión consiste en establecer si sólo existe un tipo de agente moral o si es posible manejar una visión plural de los mismos. Esta visión permitiría no identificar la máquina con el ser humano pero sí otorgarle un valor moral por encima del que poseen los objetos. Así, por ejemplo, J.P. Sullins ha sostenido que en determinadas circunstancias los robots pueden ser vistos como agentes morales, proponiendo tres requisitos: (i) autonomía respecto a programadores u operadores; (ii) su comportamiento se analiza o explica en términos de intención de hacer el bien o el mal; (iii) se comporta como responsable de otro agente moral.

No obstante también se puede seguir otro camino. Se trata del camino que han seguido algunos de los que reivindican los derechos de los animales. En efecto, en los

últimos años diversos autores y corrientes abogan por la ampliación de la agencia moral a los animales. Adela Cortina ha destacado dos planteamientos éticos en lo referido a la relación entre seres humanos-animales-derechos: el antropocentrismo (para el cual el hombre es el único ser con valor moral, el ser humano posee determinadas características que lo distinguen de los animales; los animales son instrumentos) y el biocentrismo que defiende que toda vida tiene valor moral. Así esta ampliación se hace desde este segundo planteamiento, basándose en la empatía, en las relaciones existentes o en la capacidad de sentir placer y dolor. La consideración de los animales como sujetos morales y el consiguiente reconocimiento de derechos no están reñidos con una diferenciación basada en sus distintas características.

Pues bien, esta misma argumentación podría extenderse hacia los robots, sin que ello supusiera identificarles con los humanos o con los animales.